IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 19, 2020, accepted September 4, 2020, date of publication September 9, 2020,
date of current version September 28, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3022963

ICMPv6-Based DoS and DDoS Attacks Detection Using Machine Learning Techniques, Open Challenges, and Blockchain Applicability: A Review

MOHAMMAD TAYYAB[1], BAHARI BELATON[1], AND MOHAMMED ANBAR[2], (Member, IEEE)
[1]School of Computer Sciences, Universiti Sains Malaysia (USM), Penang 11800, Malaysia
[2]National Advanced IPv6 Centre, Universiti Sains Malaysia (USM), Penang 11800, Malaysia

Corresponding author: Mohammad Tayyab (tayyab.cs.usm@student.usm.my)

This work was supported by Universiti Sains Malaysia (USM) Research University (RU) under Grant 1001/PKOMP/8014018.

ABSTRACT Although the launch of Internet Protocol version six (IPv6) addressed the issue of IPv4's address depletion, but also mandated the use of Internet Control Message Protocol version six (ICMPv6) messages in newly introduced features such as the Neighbor Discovery Protocol (NDP). This has exacerbated existing network attacks including ICMPv6-based Denial of Service (DoS) attacks and its variant form Distributed Denial of Service (DDoS) attack. Intrusion Detection Systems (IDS) aimed at tackling security issues raised by ICMPv6-based DoS and DDoS attacks have been reviewed by researchers and a general classification of existing IDSs was proposed as anomaly-based and signature-based. However, it is incredibly hard to see the overall picture of IDSs based on Machine Learning (ML) techniques with such a classification, as there is a lack of a more detailed view of the ML approach, classifiers, feature selection techniques, datasets, and different evaluation metrics. Nevertheless, recent developments in this relatively new field have not been covered such as ML-based IDSs using flow-based traffic representation. Therefore, this article specifically reviews and classifies IDSs based on ML techniques to detect ICMPv6-based DoS and DDoS attacks as single and hybrid classifiers. In addition, blockchain applicability in Collaborative IDS (CIDS) architecture based on the ensemble framework has been proposed as a solution to one of the open challenges for ICMPv6-based DoS and DDoS attacks detection problem. Moreover, this review also provides a classification of ICMPv6 vulnerabilities to DoS and DDoS attacks which would provide a reference resource for future researchers in this domain. To the best of the author's knowledge, this is the first review paper specifically focusing on IDSs based on ML techniques in this domain, as well as blockchain applicability as a possible research direction has been proposed to attract researcher's focus on building ensemble learning-based IDS models.

INDEX TERMS Intrusion detection system, CIDS, ICMPv6, DoS, DDoS, machine learning, blockchain.

I. INTRODUCTION

The global permanent deployment of IPv6 has attracted the interest of researchers to review the security issues raised by numbers of attacks and one of them is the DoS attack using ICMPv6 messages [1]. ICMPv6 has been given a vital role by the designers of IPv6 as compared to its previous version IPv4 [2]. For example, the NDP which uses ICMPv6 messages has been introduced by IPv6 as a new protocol for Stateless Address Auto Configuration (SLAAC), discovering link-layer addresses, routers discovery, and Duplicate Address Detection (DAD) processes [3]. However, these features are subjected to exploitation by the attackers to perform DoS attacks [4]. Further, many to one dimension, which is an intrinsic characteristic of a DDoS attack, is still possible in IPv6 using ICMPv6 Echo messages [5].

In literature, existing reviews on proposed IDSs for ICMPv6-based DoS and DDoS attacks detection problem has either focused on general defensive mechanisms (detection and prevention) or intrusion detection systems

The associate editor coordinating the review of this manuscript and approving it for publication was Fangfei Li.

TABLE 1. List of abbreviations used in the paper.

6LoWPAN	IPv6 over Low Power Wireless Personal Area Networks	**KNN**	K-Nearest Neighbors
AD	Anomaly-based Detection	**LWL**	Locally Weighted Learning
AH	Authentication Header	**MCC**	Matthews Correlation Coefficient
ANN	Artificial Neural Network	**ML**	Machine Learning
BN	Bayesian Networks	**MLD**	Multicast Listener Discovery
BPNN	Back Propagation Neural Network	**MLP**	Multilayer Perceptron
CART	Classification And Regression Tree	**MSE**	Mean Square Error
CIDS	Collaborative IDS	**NA**	Neighbor Advertisement
DA	Detection Accuracy	**NAv6**	National Advanced IPv6
DAD	Duplicate Address Detection	**NB**	Naive Bayes
DDoS	Distributed Denial of Service	**ND**	Neighbor Discovery
DENFIS	Dynamic Evolving Neural Fuzzy Inference System	**NDP**	Neighbor Discovery Protocol
DIO	DODAG Information Option	**NS**	Neighbor Solicitation
DIS	DODAG Information Solicitation	**P2P**	Peer to Peer
DODAG	Destination-Oriented Directed Acyclic Graph	**PCA**	Principal Component Analysis
DPI	Deep Packet Inspection	**PMTUD**	Path Maximum Transfer Unit Discovery
DR	Detection Rate	**PSO**	Particle Swarm Optimization
DT	Decision Tree	**QoS**	Quality of Service
DoS	Denial of Service	**RA**	Router Advertisement
ECM	Evolving Clustering Method	**RD**	Router Discovery
ER	Error Rate	**RF**	Random Forest
ESP	Encapsulating Security Payload	**RMSE**	Root Mean Square Error
FM	F-Measure	**RPL**	Routing Protocol for Low-Power and Lossy Networks
FN	False Negative	**RS**	Router Solicitation
FP	False Positive	**SCR**	Single Conjunctive Rule
FPR	False Positive Rate	**SD**	Signature-based Detection
ICMPv4	Internet Control Message Protocol version 4	**SDN**	Software-defined Networking
ICMPv6	Internet Control Message Protocol version 6	**SLAAC**	Stateless Address Auto Configuration
IDS	Intrusion Detection System	**SPA**	Stateful Protocol Analysis
IGR	Information Gain Ratio	**SVM**	Support Vector Machine
IHA	Intelligent Heuristic Algorithm	**TN**	True Negative
IPSec	Internet Protocol Security	**TNR**	True Negative Rate
IPv4	Internet Protocol version 4	**TP**	True Positive
IPv6	Internet Protocol version 6	**TPR**	True Positive Rate
ISP	Internet Service Provider	**USM**	Universiti Sains Malaysia
IoT	Internet of Things	**XGB**	Extreme Gradient Boosting

built on signature or anomaly-based detection approaches, as reviewed by Elejla *et al.* [6], [7], and Bdair *et al.* [8]. In general, Anomaly-based Detection (AD) builds a benign profile of network behavior by observing normal events in the network and any deviation produces an alert as a possible intrusion. On the other hand, Signature-based Detection (SD) compares patterns of known attacks with captured events by monitoring network activities to issue an alert as an intrusion.

As in [6] and [7], the focus of the reviews was on the classification of exploitation methods used to perform ICMPv6-based DoS and DDoS attacks. In addition to this, the authors proposed a taxonomy of the existing IDSs based on SD and AD approaches. Moreover, the author in [8] focused only on a brief discussion about IDSs in the detection of ICMPv6-based DoS and DDoS attacks. Furthermore, IDSs built on the flow-based representation of network traffic that was published later are missing in these reviews.

In literature, ML techniques have been reported in numbers of IDSs to detect ICMPv6-based DoS and DDoS attacks, either as single or hybrid classifiers. On the one hand, ML techniques such as Support Vector Machine (SVM), Artificial Neural Network (ANN), Decision Trees (CART, C4.5, and J48), Naive Bayes (NB), and K-Nearest Neighbors (KNN) are used in single classifier-based IDSs [9]. On the other hand, hybrid classifier-based IDSs used neuro-fuzzy techniques such as the Dynamic Evolving Neural Fuzzy Inference System (DENFIS) algorithm [10].

To the best of our knowledge, there is no review paper specifically focused on IDSs based on ML techniques for ICMPv6-based DoS and DDoS attacks detection. The ML-based intrusion detection for ICMPv6-based DoS and DDoS attacks is a relatively new field, and research in this area is gaining momentum. Therefore, the main contributions of this article are: (i) the review and classification of existing ML-based IDSs for the detection of ICMPv6-based DoS and DDoS attacks, (ii) the identification of open challenges as future research directions, (iii) proposed blockchain applicability in the ensemble framework as one of the possible solutions to these challenges, and (iv) the classification of ICMPv6 vulnerabilities that are revealed by exploitation techniques and not addressed by previous reviews.

The abbreviations used in the paper are given in Table 1. The remainder of this article is organized as follows: Section II gives a background of ICMPv6, as well as presents a classification of ICMPv6 vulnerabilities to DoS and DDoS attacks. Section III provides a brief overview of the role of ML in intrusion detection. Section IV presents the review of existing ML-based IDS models as contribution and limitations of each IDS model, as well as proposed a classification of existing ML-based IDSs detecting ICMPv6-based

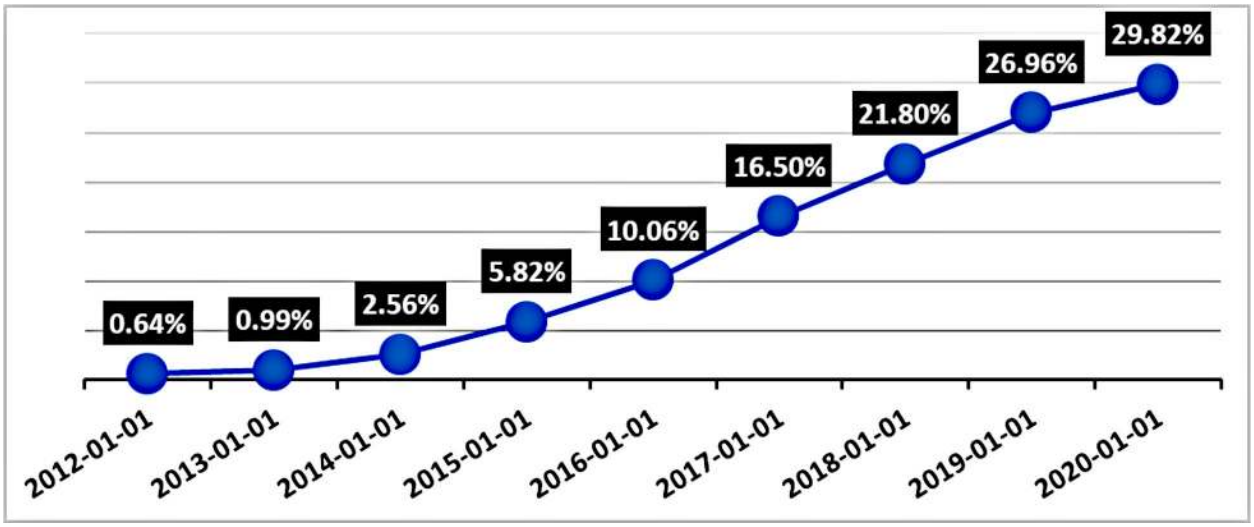

FIGURE 1. Percentage of users accessing Google via IPv6 [11].

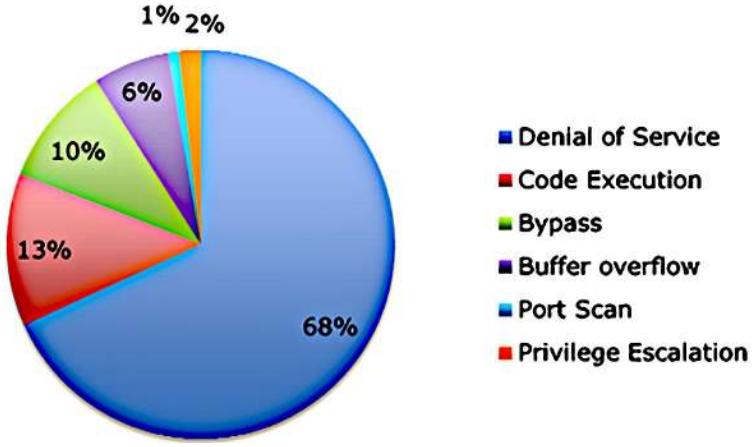

FIGURE 2. IPv6 vulnerability classes [12].

DoS and DDoS attacks. Section V lists open challenges as future research directions and Section VI discusses several recent advanced technologies as new research directions. Section VII discusses blockchain applicability as one of the possible solutions. Finally, Section VIII finishes this article with the conclusion of this work.

II. BACKGROUND

IPv4 was originally designed to provide IP addresses up to 4.29 billion [14]. Nevertheless, IP addresses provided by the IPv4 have already raised the issue of address depletion where the number of active internet users in 2019 exceeds 4.53 billion [15]. IPv6 was designed to address IPv4-related issues such as addressing depletion, security, and Quality of Service (QoS) [16]. In June 2012, the global permanent deployment of IPv6 became possible when the Internet Society, in conjunction with many large companies and organizations, held World IPv6 Launch Day [17]. According to Google statistics, Google services accessed by users over IPv6 networks reached 29.82 percent in January 2020 and the trend is still growing as shown in Fig. 1.

Although, Internet Protocol Security (IPSec) [18] and other IPv6 key features have greatly increased network security in many respects. However, Fig. 2 shows that IPv6 remains highly vulnerable to many attacks and one of them is DoS attack including ICMPv6-based DoS and DDoS attacks. The main objective of a DoS attack is to make a network or host unable to provide normal service by either attacking the bandwidth or the resources of the host [19]. The degree of a DoS attack is amplified when several coordinated devices are used to launch an attack on one or more targets to perform a DDoS attack, as shown in Fig. 3 [20]. Alternatively, DDoS attacks can also be performed remotely, as attackers can direct traffic to a vulnerable third party called a zombie using the victim's spoofed source addresses, address spoofing remains possible in both IPv6 and IPv4 [21].

As described in [RFC 4443], ICMPv6 is an extension of ICMPv4, as in addition to other new responsibilities, it still supports the same functionalities such as diagnostic, testing,

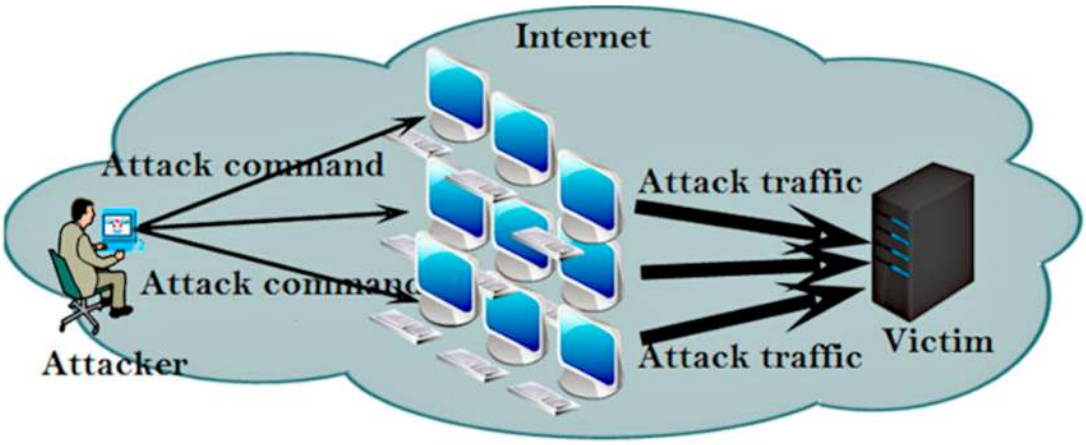

FIGURE 3. DDoS attack using coordinated devices [13].

control purposes, or to generate in response error messages for smooth communication between nodes. However, IPv6 with its new features such as SLAAC and NDP makes ICMPv6 a necessary protocol in an IPv6 network [22]. In addition, the Next Header field value of 58 in the new IPv6 header is reserved for ICMPv6 messages [2].

ICMPv6 messages are broadly categorized into two types: informational messages and error messages. The error messages (ICMPv6 message type values ranges from 1 to 127) are generated in responses to any possible errors when an IPv6 message is sent, as in the case of Path Maximum Transfer Unit Discovery (PMTUD). Whereas informational messages (ICMPv6 message type values ranges from 128 to 255) are used to exchange information between nodes, such as ICMPv6 Echo Request-Reply messages (ICMPv6 message type values 128 and 129).

However, lack of security consideration in ICMPv6 protocol functionalities makes it vulnerable to different types of attacks, including DoS and DDoS attacks [23]. The common technique used by the administrators as a solution to DoS or DDoS attacks was the blockage of ICMPv4 messages in an IPv4 network. In contrast to this, IPv6 implementation mandates the use of ICMPv6 messages for different functionalities such as PMTUD, Router Discovery (RD), and Neighbor Discovery (ND).

Nevertheless, recent research shows that malicious adversaries can misuse such messages [24]. In [25], [RFC 1981] ''Path MTU Discovery for IP version 6'' describes PMTUD that can be used to deal with fragmentation issues, such as eliminating the need for IPv6 end systems to fragment packets. On receiving a packet, if the MTU of the packets sent on a path is too large to be forwarded by some node along the path, that node will discard them and return ICMPv6 ''Packet Too Big'' message (ICMPv6 message type value 2). The PMTUD process ends when the node's estimate of the PMTU is less than or equal to the actual PMTU. This feature also supports multicast connections in which each path may have a different PMTU. Consequently, a single multicast packet may result in multiple response ''Packet Too Big'' messages from each destination and is highly vulnerable to DoS attacks [26].

The designers of IPv6 have also replaced the limited Options field in IPv4 header with the Next Header field that is more flexible and extensible as it is not part of the main header as shown in Fig. 4. Moreover, every node implementing IPv6 must support six extension headers such as Hop-by-Hop Options, Routing, Fragment, Destination Options, Authentication Header (AH), and Encapsulating Security Payload (ESP). This feature is vulnerable to DoS as by flooding of intentionally syntactically or semantically incorrect extensions invoke the ICMPv6 ''Parameter Problem'' message (ICMPv6 message type value 4) and may consume the target's resources [27].

ICMPv6 Echo messages are also vulnerable to DoS and DDoS attacks. A malicious node or a group of coordinated devices can send large amounts of ICMPv6 Echo-Request messages (ICMPv6 message type value 128) to the victim with their source targeting at another IPv6 node or an invalid IPv6 address. This can waste the resources of the victim by receiving ICMPv6 Echo-Reply messages (ICMPv6 message type value 129) and can cause it to stop responding to other requests.

The Multicast Listener Discovery (MLD) protocol enables MLD-capable routers to maintain node information listening to multicast addresses [28], hence allowing the forwarding of packets destined for these addresses. To keep track of the multicast addresses that have listeners, a query router that is capable of maintaining this information regularly sends general query messages (ICMPv6 message type value 130) to the link-scope all-node address (FF02::1). Then, the listening nodes respond to the multicast address being reported with a report message (ICMPv6 message type value 131). However, a malicious node can abort this forwarding of multicast-destined packets by sending a spoofed done (ICMPv6 message type value 132) message [29]. In contrast to this, MLD general query messages flooding aiming to target a specific multicast group would compromise all multicast group listeners. For example, flooding of the MLD

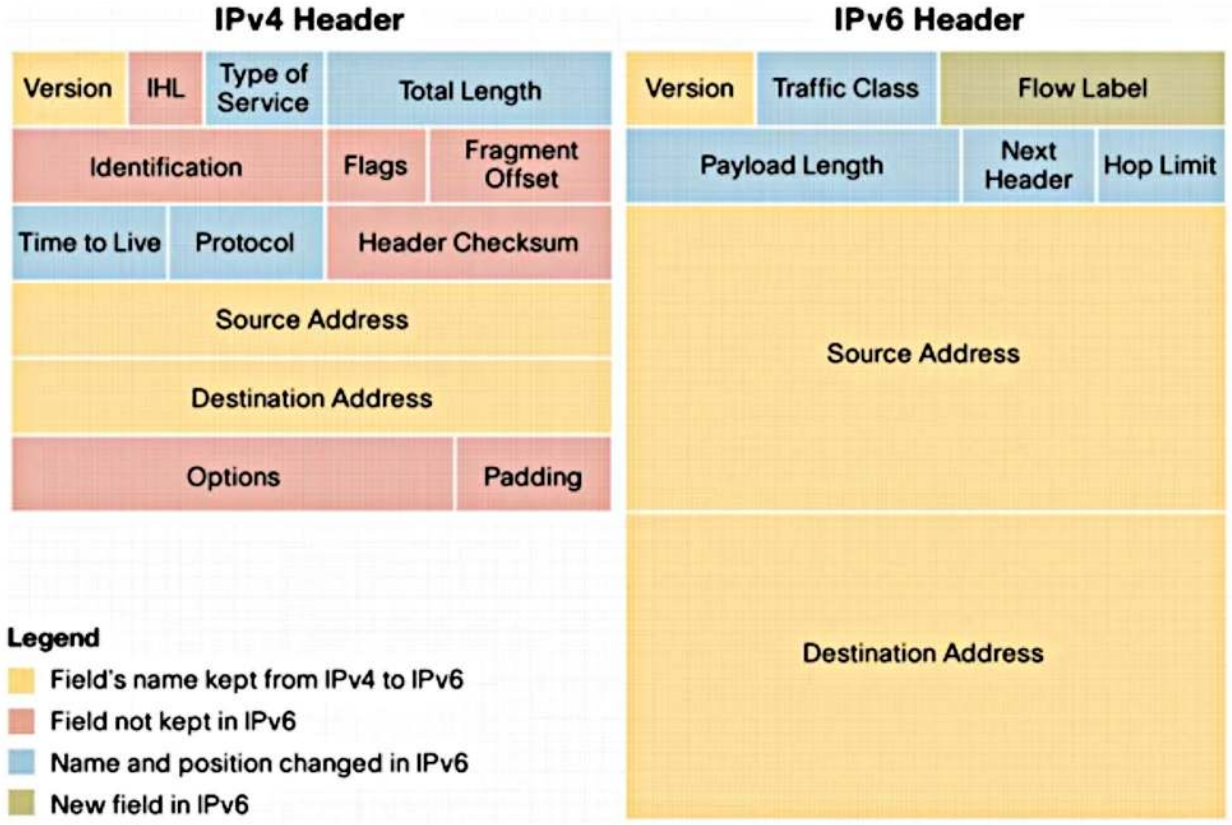

FIGURE 4. Comparison of IPv4 and IPv6 headers [12].

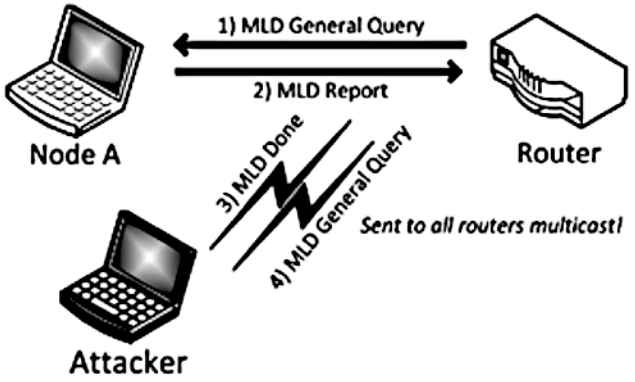

FIGURE 5. MLD exploitation [30].

general query message targeting the link-scope all-routers address (FF02::2), inevitably results in all routers on the link listening to FF02::2; thus, these routers are compromised by the flooded traffic, as shown in Fig. 5.

NDP primarily uses ICMPv6 messages (ICMPv6 message type values ranges from 133 to 137) for features such as ND and SLAAC [3]. Trust vulnerability in NDP messages allows DoS attacks to be performed by exploiting the functions of NDP protocol [31]. For example, the flooding of Neighbor Solicitation (NS) messages (ICMPv6 message type value 135) and Neighbor Advertisement (NA) messages (ICMPv6 message type value 136) would eventually consume a portion of the victim's resources or may cause a complete stoppage of its services while it tries to respond to the malicious messages [32]. In the same fashion, Router Solicitation (RS) messages (ICMPv6 message type value 133) and Router Advertisement (RA) messages (ICMPv6 message type value 134) flooding and malformed RA prefix advertisements can cause traffic interruption and routing storms. In addition, redirect vulnerabilities can be exposed by Redirect messages (ICMPv6 message type value 137) that may redirect hosts on a link to various off-link routers which can cause route flapping (DoS).

In [33], DoS and DDoS attacks are described separately. First DoS attacks are classified into five categories which include (1) network device level, (2) operating system level, (3) application level, (4) data flood, and (5) protocol feature attack. Whereas the classification of DDoS attacks consists of two levels. At the first level, attacks are classified according to their degree of automation, exploited vulnerability, attack rate dynamics, and their impact. Specific characteristics of each first-level category are recognized in the second level. In continuation of this, the authors [6] proposed a classification of ICMPv6-based DoS and DDoS attacks based on exploitation techniques such as flooding-based, amplification-based, and NDP exploitation-based, as shown in Fig. 6. In addition, Swami *et al.* [34] discussed common types of DDoS attacks and presented a taxonomy of DDoS attacks as shown in Fig. 7.

Traditionally, the primary reason for ICMPv6-based DoS and DDoS threats is ICMPv6 messages-based vulnerabilities that are exposed through the exploitation techniques.

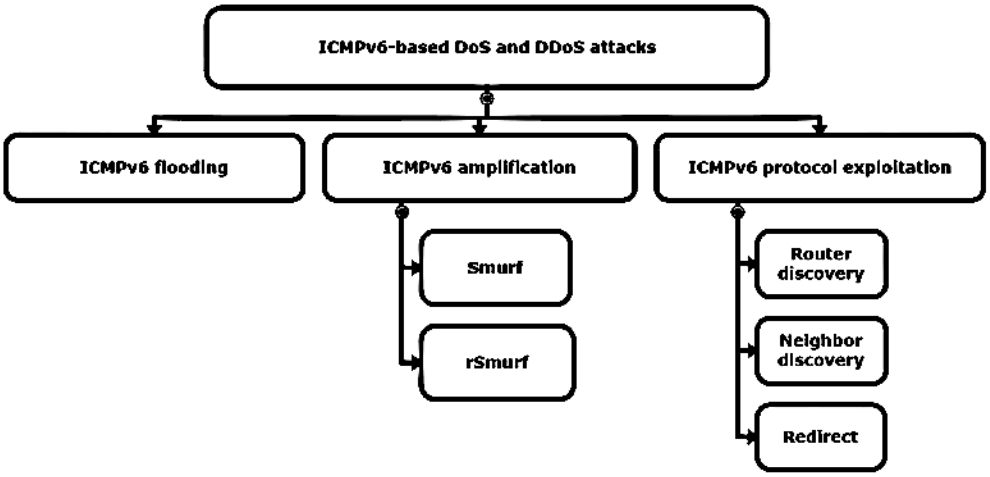

FIGURE 6. Categorization of ICMPv6-based DoS and DDoS attacks [6].

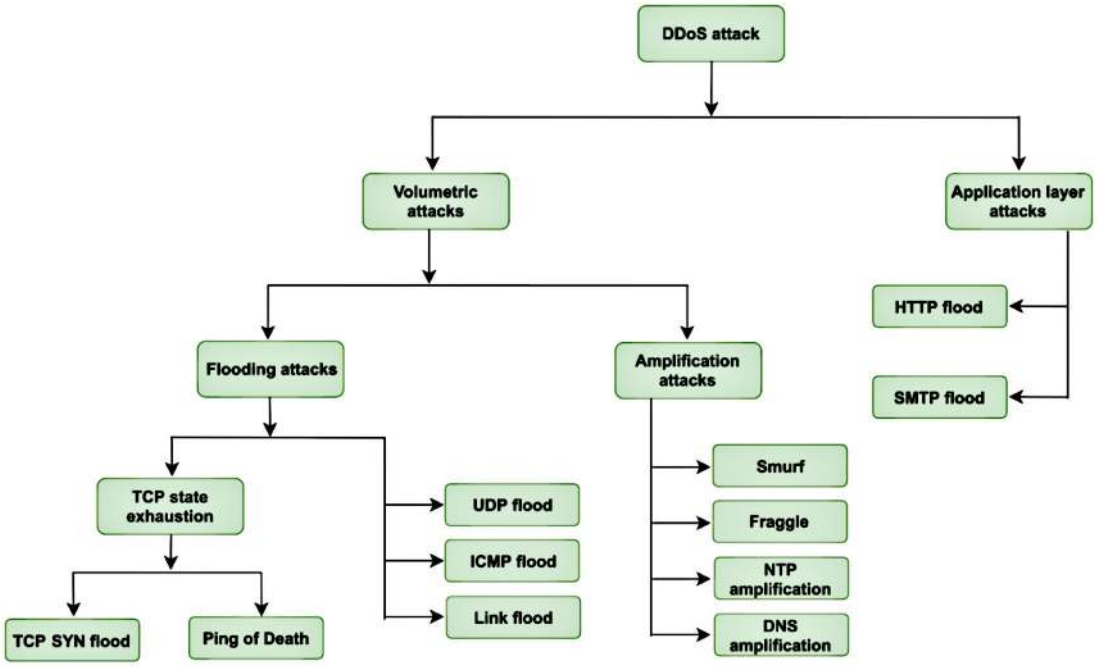

FIGURE 7. Taxonomy of DDoS attacks [34].

The intrusion detection problem needs the researcher's focus on ICMPv6 messages-based vulnerabilities for DoS and DDoS attacks rather than on the exploitation techniques. As a contribution, we, therefore, propose a new classification of ICMPv6 messages-based vulnerabilities for DoS and DDoS attacks which includes two levels. First, the ICMPv6 vulnerabilities are classified as ICMPv6 error messages-based, ICMPv6 informational messages-based, and ICMPv6 subprotocol messages-based. In the second level, vulnerabilities exposed through exploitation techniques in different protocol features are mentioned. Therefore, ICMPv6 vulnerability classes for DoS and DDoS attacks can be categorized as shown in Fig. 8.

Moreover, the importance of ICMPv6 for the smooth functioning of an IPv6 network is of prime interest. In other words, IPv6 key features cannot operate without the support of the ICMPv6 protocol [35]. Therefore, the only way to prevent ICMPv6 vulnerabilities from exposure to DoS and DDoS attacks is to deploy a network-wide detection system for continuous monitoring as well as to identify any malicious behaviors of ICMPv6 messages that could lead to such an attack.

III. MACHINE LEARNING ROLE IN INTRUSION DETECTION

This section gives a brief about ML as an evolutionary field of artificial intelligence where pattern recognition in data is achieved through efficient adaptive methods. These methods allow a machine to adapt accordingly by building a model from example inputs to make data-driven predictions or decisions, rather than following strictly static

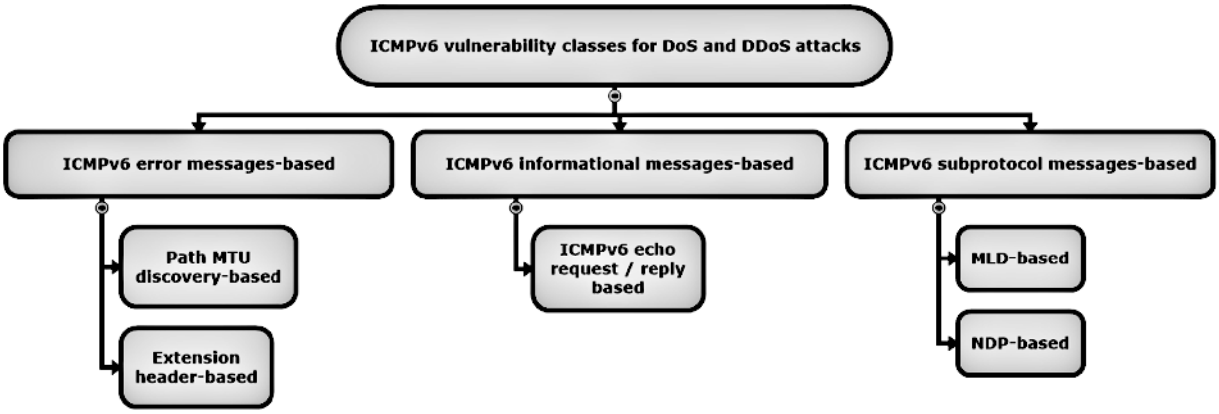

FIGURE 8. Classification of ICMPv6 vulnerabilities to DoS and DDoS attacks.

program instructions. The role of ML in the intrusion detection field is to build a model for a multinomial classifier problem that can classify network events as normal or attack events, as in the case of DoS and DDoS attacks. Recent research shows that, compared to traditional IDS solutions, researchers have shifted their focus towards developing ML-based IDS solutions [36]. In the literature, ML-based IDS models achieved interesting results; from 86.53% [37] to over 99% [38] in detection accuracy and a significant decrease in false-positive from around 4% [39] to 0.01% [40].

In addition, Verma and Ranga [41] performed a thorough performance analysis to find substantial variations between various ML techniques using statistical analysis methods such as Friedman and Nemenyi tests. The authors stated that if the application requires high accuracy and low false-positive rate, then ML techniques such as CART, Multilayer Perceptron (MLP) are useful. While Random Forest (RF) and Extreme Gradient Boosting (XGB) have proven their usefulness if time is more crucial.

A. CLASSIFIER TYPES

The classification approach using ML techniques can be categorized as single, ensemble, or hybrid depending on the number and the way in which different techniques work together to solve a problem [42].

1) SINGLE

IDS model under this category can be designed by utilizing only one technique such as clustering, classification, or association. In recent research, classifiers such as SVM, NN, Decision Trees (CART, C 4.5, and J48), NB, and KNN have been used to design single classifier-based IDS models.

2) HYBRID

In this category, an IDS model typically combines two or more functional components with the intuition to improve performance as an advantage over a single classifier approach. Classifiers implementation in this approach works in two stages, first one aims at optimizing the learning performance (i.e. parameter tuning). The second stage, then use intermediate results to predict the final output. The design of a hybrid classifier can be based on cascading different classifiers, such as the DENFIS algorithm [10].

3) ENSEMBLE

Ensemble models add another dimension to achieve performance intuitive benefits by combining the opinions of multiple learners. Ensemble methods, with access to multiple processors, are the ideal choice for training and testing time efficiency because they are inherently parallel in nature. The ensemble model's implementation can be accomplished in two ways, one is training multiple classifiers on the same dataset and the other is training a single classifier on multiple datasets. After the training phase, the data item is assigned to the class to which the majority of classifiers point at the time of testing.

B. LEARNING TECHNIQUES

In the literature, ML techniques are generally classified as supervised or unsupervised depending on the presence and absence of the labeled data, and what we are trying to predict from the dataset [42]. Supervised learning uses the training data to create a function, in which each of the training data contains a pair of the input vector and output (class label). Classifier's training is achieved by computing the approximate distance between the input-output examples to create a classifier (model). Once the trained model is ready, it can classify unknown examples into a learned class label.

In contrast to supervised learning, unsupervised learning take on an unlabeled dataset and assign the items to certain groups. Hyper-parameter tuning in this technique refers to the number of clusters to be labeled within the dataset. Unsupervised learning's main aim is to uncover the hidden groups in the dataset, rather than predicting the label for some data items.

C. FEATURE SELECTION TECHNIQUES

Features selection techniques help in building an effective learning method either by choosing a subset of significant features or evacuating the insignificant and excess features from the dataset. The datasets with a large number of features for the ML-based IDS model are not feasible and may result

in reduced performance [43]. In literature, researchers have been using many feature selection algorithms such as Information Gain Ratio (IGR) and Particle Swarm Optimization (PSO). In addition to this, dimensionality reduction technique such as Principal Component Analysis (PCA) has also been reported to resolve the issue of data dimensionality during the feature extraction phase, which in turn minimizes the entire volume of training data that needs to be processed; thus, reduces the computation time to yield a more precise classification.

D. DATASET

The intrusion dataset containing both normal and malicious network traffic is the essential component that can be used for benchmarking the performance of an IDS. Benchmarking refers to the performance of an IDS using performance evaluation metrics results obtained during experiments after applying various data mining techniques [44]. The intrusion datasets are mostly generated in the following two ways:

- A testbed laboratory or virtual environment is set up to simulate the different network topologies. In this setup, scripts are used to generate artificial attack traffic and traffic samples are collected from the network. Ease of implementation encourages to generate this type of dataset, where all attack types can be manually injected. The major drawback of such datasets is that they do not represent the real-world network traffic scenarios. There is a high probability that intrusion detection systems evaluated on such datasets are not guaranteed to give similar results in a real-world deployment.
- The collection of network traffic from real-world networks is another way for intrusion detection dataset generation as these datasets represent the actual nature of network traffic. However, there is a possibility that these datasets may not contain all the required types of attacks. Although the realistic nature of these datasets gives a true real-world representation, some obstacles make it difficult to build. For instance, an organization's privacy concern limits the collection of traffic samples from their network. In addition, legal laws also bound them to not allow publishing of actual data in the public domain.

E. PERFORMANCE EVALUATION METRICS

Cross-validation and supplied test set approaches have been applied to evaluate the performance of an IDS by obtaining the results in the form of different performance evaluation metrics [45]. Experimental results using these metrics have been used by researchers to compare their results with already existing approaches.

True Positive (TP): An attack traffic instance correctly classified as belonging to attack class.

False Positive (FP): A normal traffic instance incorrectly classified as belonging to attack class.

True Negative (TN): A normal traffic instance correctly classified as a normal class instance.

False Negative (FN): An attack traffic instance incorrectly classified as a normal class instance.

Detection Accuracy (DA): Detection accuracy measures the ratio of correct predictions over the total number of instances evaluated.

$$DA = \frac{TP + TN}{(TP + FP + TN + FN)} \tag{1}$$

Error Rate (ER): Also referred to as misclassification error, measures the ratio of incorrect predictions over the total number of instances evaluated.

$$ER = \frac{FP + FN}{(TP + FP + TN + FN)} \tag{2}$$

True Positive Rate (TPR): The intrusions which are correctly classified as an attack are also known as sensitivity.

$$TPR = \frac{TP}{(TP + FN)} \tag{3}$$

False Positive Rate (FPR): Often referred to as false alarm. These are the normal patterns that were incorrectly classified as an attack.

$$FPR = \frac{FP}{(FP + TN)} \tag{4}$$

True Negative Rate (TNR): The normal patterns that were correctly predicted as normal are also known as specificity.

$$TNR = \frac{TN}{(TN + FP)} \tag{5}$$

Precision (P): The positive patterns that are correctly predicted from the total predicted patterns in a positive class.

$$P = \frac{TP}{(TP + FP)} \tag{6}$$

Recall (R): The fraction of positive patterns that are correctly classified.

$$R = \frac{TP}{(TP + TN)} \tag{7}$$

F-Measure (FM): This metric represents the harmonic mean between recall and precision values.

$$FM = \frac{2 * P * R}{(P + R)} \tag{8}$$

Matthews Correlation Coefficient (MCC): This metric measures the correlation between the predicted results and the real data.

$$MCC = \frac{(TP.TN) - (FP.FN)}{\sqrt{(TP + FP).(TP + FN).(TN + FP).(TN + FN)}} \tag{9}$$

IV. EXISTING ML-BASED IDS MODELS DETECTING ICMPV6-BASED DOS AND DDOS ATTACKS

Traditionally, the approaches used for intrusion detection are broadly classified as three categories: SD, AD, and Stateful Protocol Analysis (SPA) [46]. Their conceptual descriptions are as follows: SD uses a packet header and payload to compare signatures (known attacks) against captured events

TABLE 2. Comparison of intrusion detection approaches.

Signature-based	Anomaly-based	SPA-based
Pros		
• Highly effective and less complex approach to detect known attacks. • Produces low false alarms through detailed contextual analysis.	• Efficient in the detection of zero-day vulnerabilities. • Facilitates the detection of privilege abuse by building benign network behavior profiles.	• Monitors known protocol states rigorously. • Traces illegitimate sequence of activities to distinguish normal or malicious behavior.
Cons		
• Almost unable to detect zero-day vulnerabilities and variations in known attacks. • Adds overhead to keep the signature database updated all the time.	• Relatively produces more false alarms due to the dynamic nature of observed events. • The nature of the events being constantly changed makes it almost unavailable during the rebuilding of benign profiles.	• Consumes a good amount of monitoring node resources for protocol state tracing and examination. • Ineffective to inspect attacks outside the scope of benign protocol behaviors.

for the detection of potential intrusions by monitoring network events. In AD, benign profiles representing normal behavior are derived by monitoring regular activities in the network and any deviation (anomaly) from normal behavior is considered a possible intrusion. Profiles can be either static or dynamic containing many attributes (features) and are developed over time. Nevertheless, SPA focuses on complete semantics of protocols as mentioned in the specification and any out of range value is considered an intrusion. Each of these categories has its pros and cons in intrusion detection as shown in Table 2.

In [47], the authors mentioned two major detection methodologies as Deep Packet Inspection (DPI) and SPA to detect attacks in the network traffic. In addition, due to the limitations mentioned in Table 2, the authors also proposed a taxonomy for intrusion detection systems based on the method used for attack detection as three subclasses: Statistics-based, ML-based, and others for intrusion detection.

Moreover, considering the importance of DoS and DDoS threats using ICMPv6 messages in the IPv6 network, several other classifications have been proposed by researchers such as [6], [7], and [8]. Based on the ML approaches, existing ML-based IDS models can be categorized into two classes as single or hybrid for the detection of ICMPv6-based DoS and DDoS attacks. Furthermore, a single classifier-based IDS models can be subdivided into two classes as IDS models based on packet-based features and flow-based features. Fig. 9 shows the classification of existing ML-based IDS models that can detect ICMPv6-based DoS and DDoS attacks. This section presents a review and limitations of existing IDS models based on ML techniques for the detection of ICMPv6-based DoS and DDoS attacks.

A. HYBRID CLASSIFIER-BASED IDS

ML was first applied for the detection of NDP-based DoS attacks by Yao *et al.* [48]. Inspired by the artificial immune system theory introduced by Forrest *et al.* [49], that detects the self (normal) or nonself (anomaly) in an immune system through a network security system, the authors proposed a fuzzy logic-based anomaly detection algorithm. The genetic algorithm is used in self-space (normal) to generate fuzzy rules that cover anomalies in network data based on the fitter function as a distance from the fuzzy parameter (self-space radius). Features extracted from network data have been defined as the input vector, and the distance calculated through the fitter function determines the vector as normal or anomaly.

Real network data from the CERNET2 national center, which has an IPv6 setup at Tsinghua University in China, was used to train the proposed model for evaluation purposes. Detection Rate (DR) and FPR were used as the performance metrics which shows that the proposed model achieved DR of 85% along with FPR of 2% when the value of the fuzzy parameter was less than 0.6. Despite the optimum results shown in the experiments, details such as features selection and their ranking algorithm is missing in the paper. Further, the insertion of attack traffic with real data for testing is highly susceptible to produce a biased dataset [7].

Saad *et al.* [50] applied another hybrid classification-based learning technique to detect DDoS attacks based on the ICMPv6 Echo-Request flood. The proposed IDS model uses the Evolving Clustering Method (ECM) [51] for online clustering, followed by the DENFIS algorithm for normal or attack traffic learning. The generation of new fuzzy rules depends on the creation of a new cluster produced through ECM otherwise one or more existing fuzzy rules are updated. By applying the back-propagation method, fuzzy rules are optimized for every prediction and dynamically selected to predict network traffic as normal or malicious.

The evaluation of the model was performed at the National Advanced IPv6 (NAv6) center in Universiti Sains Malaysia (USM) after setting up a testbed network of 6 hosts. In the testbed network, data sets for training and testing were prepared using a self-developed program containing a flood of 1000 to 1500 ICMPv6 Echo-Request packets for the attack while normal traffic of 10 to 15 packets. Performance evaluation metrics such as DA and Root Mean Square Error (RMSE) are used to evaluate the detection performance of the model using a testing dataset of 400 samples. The result of experiments shows that with 0.26 RMSE, the model offered a high DA of 98.3 percent. However, the key drawback of the

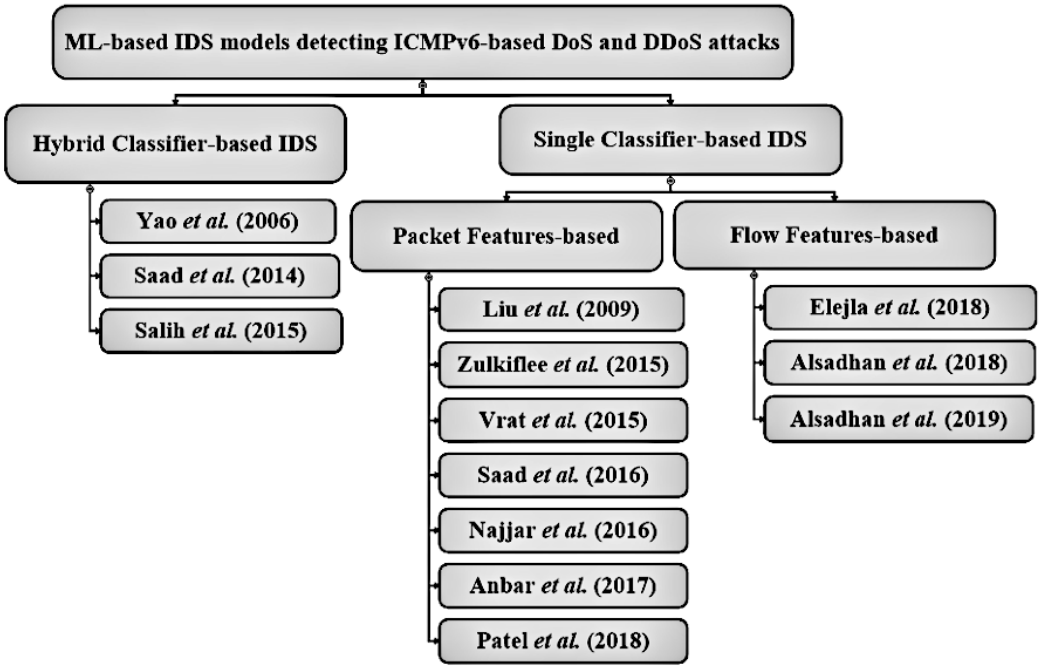

FIGURE 9. Classification of existing ML-based IDS models detecting ICMPv6-based DoS and DDoS attacks.

model is the sole detection of the ICMPv6 Echo-Request flooding attack [52]. Moreover, the size of the datasets of 2000 ICMPv6 Echo-Request packets cannot reflect the actual behavior of the DDoS attack [53].

Salih *et al.* [54] proposed another classification-based model to predict ICMPv6-based covert channels in an IPv6 network. The authors have mentioned a scenario where a covert channel can be established by the malformed data portion of an ICMPv6 Echo-Request and ICMPv6 Echo-Reply packet. The model starts by inspecting the packet's header to classify pattern behavior by identifying insignificant or null values in the data portion of the ICMPv6 packet. The proposed model follows a framework to generate a training dataset by applying Intelligent Heuristic Algorithm (IHA) along with the C4.5 classifier. Then, the NB classifier is used to predict the class as normal or covert after analyzing instances that include values of type, code, and payload fields of an ICMPv6 packet in the dataset.

Two different datasets, one real (DARPA 1999) and another generic were prepared for experiments done in two phases. In the first phase, the author used 10-fold cross-validation techniques on a simulated dataset containing 11 features with the DA of 94.47%. Whereas the second phase is used to improve the proficiency of the model on the DARPA 1999 dataset with a DA of 96.55%. However, the proposed model is limited to detect covert channels using ICMPv6 packets and does not include DoS and DDoS attacks [52].

B. SINGLE CLASSIFIER-BASED IDS

1) PACKET FEATURES-BASED IDSs

In 2009, Liu and Lai [55] used the association rule mining approach by an improvement to the Apriori algorithm [56]. The authors have mainly targeted DDoS attacks based on ICMPv6 Echo-Request flood and DoS attacks by malformed router header. To recognize an attack or normal traffic, the proposed model preprocesses network data by representing a transaction record as a series of connection attempts using ICMPv6-Echo messages to different ports on the victim host within a given duration threshold. Training and testing datasets consisting of 5000 connection records were prepared after extracting six features with an addition of class label as normal or abnormal for association mining.

Eight association rules were derived with minimum support and confidence of 0.1 and 1 respectively. Experimental results show that DA of 72% was achieved with minimum support of 0.1 for normal connection records. Nevertheless, a small network of only four PCs used for evaluation purpose lacks DDoS attacks behavior along with low DA of 72% [7]. Further, nonqualified features such as port, flag, and state are used to detect ICMPv6-based DDoS attacks.

The SVM was first introduced by Zulkiflee *et al.* [57] as a classification technique to detect ICMPv6-based DoS attacks. PSO proposed by Moraglio *et al.* [58] is used as a method for feature selection to produce optimum DA. The authors have mainly focused on a framework consisting of five phases to identify features that have contributed to the enhancement of the proposed model's detection capability. Five features identified through the framework are TimeIntvl, SrcIP, SrcPort, DstPort, and Protocol.

Average DA of 99.5% reflected the capability to differentiate between normal and attack data using the datasets containing 250,008 records. Though the authors have claimed high DA, the proposed technique is only limited to flooding-based DoS attacks using RA messages of NDP [44]. In addition, less

detail about the experimental environment and attack tools is given.

Another model for DoS attacks including ping-of-death, smurf, and teardrop attacks in an IPv4 or IPv6 network was proposed by Vrat *et al.* [59]. The authors have used a modified version of the KDD Cup 1999 [60] dataset using 11 features for training and testing of the proposed model. NB, Decision Table, J48, and PART classifiers are applied to analyze and predict the DoS anomalies existing in the given dataset.

Performance evaluation metrics such as precision, TPR, FPR, F-measure, and DA are used to evaluate the detection performance of the model. J48 classifier produced the best results with 97% DA using the original KDD Cup 1999 dataset with 41 features and a modified dataset with 11 features. However, the use of an offline dataset prepared under a small network topology lacks a real-world DDoS attack scenario [59]. In addition, the detection is limited to DDoS attacks using ICMPv6 Echo-Request messages [52].

In continuation of the author's previous work, the Back Propagation Neural Network (BPNN) algorithm was again applied by Saad *et al.* [61] to classify flooding-based DDoS attacks using ICMPv6 Echo-Request messages. Feature selection was performed using the IGR algorithm for features ranking, and PCA to reduce the dimensionality of the data as required for feature set extraction. Features such as the number of ICMPv6 packets, source IP, and destination IP addresses are aggregated and used as input for threshold-based activation function in BPNN to detect ICMPv6-based flooding behavior.

The performance of the proposed model was evaluated using a real dataset generated at NAv6 in USM. Experimental results have recorded up to 98.3% DA at a threshold value of 1000 packets/second using ICMPv6 Echo-Request messages. Further, the authors have claimed that the Mean Square Error (MSE) value of 0.00030683 at epoch 8 is achieved by increasing the volume and time of training of the BPNN model. The model's detection capability is limited to flooding-based DDoS attacks using ICMPv6 Echo-Request messages. Further, nonqualified features such as time would lead to misclassification [52].

Najjar *et al.* [62] used a strict anomaly detection approach introduced by Sasha and Beetle [63] to detect violations in normal RS and NS packets flow. A feature set containing eight features is used to distinguish between normal or anomaly flow against the protocol constant values defined in [RFC 4861 and RFC 4862] for NDP messages. Dataset used in the experiments was generated from a small virtual network and attack traffic was injected by the Hacker's Choice (THC) IPv6 attacking tool [64].

Dataset preprocessing is accomplished by eliminating 1129 duplicate instances of the total 1639 instances in the dataset. Six ML classifiers using Weka [65] data mining software are used to classify 510 instances as normal or attack flows. Research outcome shows that the C4.5 classifier outperformed with a DA of 100% for RS and NS flooding attacks. In spite of such a high DA, the dataset used in the proposed model only contains two DoS attack scenarios based on RS and NS messages [64].

The Detection of RA flooding attack using SVM [66] was proposed by Anbar *et al.* [67]. Ranking of the feature set is accomplished using the IGR algorithm whereas PCA was applied to reduce data dimensionality in the proposed model. Two separate real datasets, each containing 199138 instances, are used for training purposes to predict RA flooding attack. The first training dataset has 9 features whereas the second training dataset with 5 features was prepared after the feature reduction phase.

For performance evaluation, a testbed network topology consisting of 6 nodes is used at NAv6 in USM [68]. The validity of the features, selected after reduction, was proved with 98.55% DA and FPR of 3.3%. Whereas the original feature set consisting of 9 features recorded 94.93% DA along with FPR of 4.2%. However, the scope of the detection was limited to DoS attacks using RA messages [52].

The network traffic classification approach for ICMPv6 packets, a key technique for the mitigation of DDoS attacks by Internet Service Providers (ISP), is employed by Patel *et al.* [69]. This research extracted four traffic features that were used during the classification phase namely source IP, destination IP, interval, and duration. A sampling of collected network data is accomplished by adding the Protocol as a class label to achieve the aim of classifying the unknown network data at the testing stage of the model.

To achieve optimal performance, a comparative performance analysis is performed with four ML classifiers such as NB, KNN, DT, and SVM. Cross-validation performance analysis revealed that KNN obtains a minimum variance of 0.0000084 for 98.3% class prediction accuracy over different runs using the Python programming language for k=11. However, network traffic classification for ICMPv6 subprotocol, as in the case of NDP-based DDoS attacks, is outside the scope of this model.

2) FLOW FEATURES-BASED IDSs

A flow-based approach for network traffic representation was first introduced by Elejla *et al.* [70] to detect ICMPv6-based DDoS attacks. Flow is defined as a series of ICMPv6 packets of the same ICMPv6 type that are sent from one IPv6 source to another IPv6 destination within a given time interval. Eleven features are extracted to generate training and testing datasets that are required to build the model using seven classifiers.

Cross-validation and supplied test set approaches of data mining are applied to build the model using seven classifiers such as C4.5, SVM, NB, KNN, NN, RF, and Single Conjunctive Rule (SCR). DA and FPR are used as performance evaluation metrics to advocate a flow-based approach for an efficient IDS over packet-based representation. Experimental results using 10-fold cross-validation and supplied test set approaches reveal almost the same DA and FPR as 85.67% and 17.1% respectively by the C4.5 and SVM classifiers. High FPR is pointing to

TABLE 3. Comparison of existing ML-based IDS models detecting ICMPv6-based DoS and DDoS attacks.

IDS Model Reference No.	Classifier(s)	Dataset(s)	Feature Selection Technique(s)	Metric(s)	Result
Hybrid Classifiers-based					
[48]	Fuzzy Logic with GA	CERNET2	None	DR and FPR	DR-85% and FPR-2%
[50]	DENFIS	Custom	ECM	DA and RSME	DA-98.3% and RSME-0.26
[54]	IHA, NB, and DT (C4.5)	10%-DARPA 1999 and Custom	None	DA, TPR, FPR, and Precision	DA (DARPA-96.55% and Custom-94.47%)
Single Classifiers-based					
[55]	Apriori	Custom	None	DA	DA-72%
[57]	SVM	Custom	PSO	DA	DA-99.95%
[59]	NB, DT, J48, and PART	KDD Cup 1999	None	DA, TPR, FPR, Precision, and F-measure	DA-97% with J48
[61]	BPNN	Real traffic and injected attacks	IGR and PCA	DA and RSME	DA-98.3% and RSME-0.00030683
[62]	ZeroR, OneR, NB, C4.5, KNN, and SVM	Custom	None	DA	DA-100% with C4.5
[67]	SVM	Real traffic and injected attacks	IGR and PCA	DA, FPR, Precision, and Recall	DA-98.55% and FPR-3.3%
[69]	NB, KNN, DT, and SVM	Real traffic	None	Class Prediction	98.30%
[70]	C4.5, SVM, NB, KNN, NN, RFT, and SCR	Real traffic and injected attacks	None	DA and FPR	DA-85.67% and FPR-17.1% with C4.5
[71]	DT, NB, RFT, and MLP	Real traffic and injected attacks	IGR and PCA	DA, FPR, and TPR	DA-87.3% and FPR-0.25 with DT
[72]	BN, DT, and NB	Real traffic and injected attacks	IGR	DR, FPR, and TPR	DR-96.46% and FPR-0.0004 with BN

the biased selection of features without feature ranking algorithms [52].

Alsadhan *et al.* [71] have also proposed an NDP-based DDoS attack detection model using the flow-based approach for network traffic representation. Flow in the proposed model is defined as a stream of NDP messages of the same NDP type from one IPv6 source to another IPv6 destination address over a threshold period. PCA is used for constructing a predictive model to ensure the contribution of extracted features in detecting NDP-based DDoS attacks in a five-phase framework.

DT, NB, RF, and MLP classifiers are used to evaluate the performance of the model in the detection of NDP-based DoS attacks. DT performed well with 84.3% DA among other classifiers using a dataset with 12 features without ranking features. In addition, the implementation of PCA with DT using a different dataset with 10 features has further improved DA to 87.3%. However, this proposed model does not simulate an online scenario as it has used a virtual testbed network to generate attack traffic.

Alsadhan *et al.* [72] have again utilized the flow-based representation of network traffic to propose a model based on Locally Weighted Learning (LWL) for the detection of NDP-based DDoS and replayed attacks. Englert [73] introduced the LWL technique where a local function can leverage the current point of interest prediction using a subset of training data, regardless of a global function that needs complete training data for prediction. The model has employed the LWL technique using Bayesian Networks (BN), DT and NB classifiers as base learners.

The dataset used has seven features to train the model namely ICMPv6Type, PacketsNumber, Destination, ByteNumber, SAT, BytesRatio, and MClass. Comparative analysis of all three base learners revealed that the LWL-BN model took the lead by producing a higher DA of 96.48% with 0.0004% FPR. Although the model has produced high DA, attack traffic generated from a single source in a simulated network lacks a real-time DDoS environment. In addition, feature set selection was based on domain knowledge that lacks a solution to the data dimensionality issue that exists in the dataset.

C. SUMMARY

This review presents the current state of the art of ML-based IDSs for the detection of ICMPv6-based DoS and DDoS attacks. Based on the classifier's design we present a classification of existing ML-based IDSs into single and hybrid categories for the detection of ICMPv6-based DoS and DDoS attacks. Table 3 gives an overall idea about the ML approach, classifier used, and feature selection techniques along with the result produced by each model. Although existing ML-based IDS models such as [50] and [61] have promised high DA in the detection of ICMPv6 Echo messages-based DDoS anomalies, the growing scale of the problem with respect to the analysis of traffic produced by DDoS attacks can have a significant impact on the performance of ML-based IDSs using sequential hardware. Therefore, IDS models based on ML techniques is still in its infancy stage for the detection of ICMPv6-based DoS and DDoS attacks.

TABLE 4. Scope of detection of the existing ML-based IDS models detecting ICMPv6-based DoS and DDoS attacks.

IDS Model Reference No.	MTU Discovery-based Attacks	Extension Header-based Attacks	ICMPv6 Echo messages-based Attacks	MLD-based Attacks	NDP-based Attacks
Hybrid Classifiers-based					
[48]	No	No	No	No	Yes
[50]	No	No	Yes	No	No
[54]	No	No	Yes	No	No
Single Classifiers-based					
[55]	No	Yes	Yes	No	No
[57]	No	No	No	No	Yes
[59]	No	No	Yes	No	No
[61]	No	No	Yes	No	No
[62]	No	No	No	No	Yes
[67]	No	No	No	No	Yes
[69]	No	No	Yes	No	No
[70]	No	No	Yes	Yes	Yes
[71]	No	No	No	No	Yes
[72]	No	No	No	No	Yes

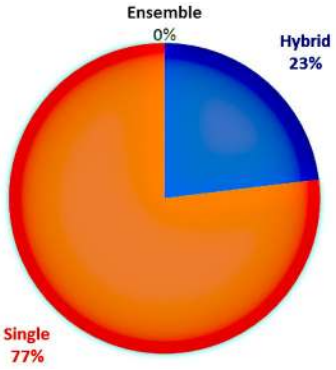

FIGURE 10. Types of classifiers used in ML-based IDS models detecting ICMPv6-based DoS and DDoS attacks.

In addition, the classifier's ability to learn and build the model mainly depends upon the feature set and the solution to class imbalance factors existing in the dataset. The models such as [48], [62], and [70] lack features selection schemes needed to validate the relevance of the features in generating the datasets. Table 4 gives an insight into the scope of detection that is covered by each ML-based IDS model for ICMPv6-based DoS and DDoS attacks. Some researchers as in [54] and [59] have focused only on DDoS attacks using ICMPv6 Echo Request-Reply messages while others such as [71] and [72] targeted only NDP-based DoS attacks.

The classifier type as single, hybrid, or ensemble plays a significant role in ML-based IDS in enhancing the ability of the classifiers to predict possible anomalies that exist in the dataset. Fig. 10 shows existing ML-based IDSs followed only two approaches namely single and hybrid. The potential of an ensemble framework using a parallel/distributed environment can be highly effective in the detection of DDoS attacks based on ICMPv6 messages, hence this area of research also needs focus from researchers.

V. OPEN CHALLENGES

This review gives an overall picture of the existing ML-based IDS models for the detection of ICMPv6-based DoS and DDoS attacks. The contribution of this work may be used as a source for future research directions in this area. Most notably, IDS models based on ML techniques are essential requirements for achieving higher accuracy in the detection of ICMPv6-based DoS and DDoS attacks; thus, we have identified several open challenges for research in this domain from our comprehensive review.

A. PERFORMANCE BOTTLENECK IN REAL-TIME

In most of the existing ML-based IDS models, the offline networking environment is used for training and testing of these models for the detection of ICMPv6-based DDoS attacks. ML-based IDS models trained in such an offline environment may not achieve high accuracy when detecting real ICMPv6-based DDoS attacks. Specifically, the ICMPv6-based DDoS attack may affect the performance of a standalone ML-based IDS by an increase in the volume of data to be processed in real-time. Hence, existing ML-based IDS models may not work well in real-time. Therefore, there is an ever-increasing need to improve the accuracy in real-time by providing an IDS model based on ensemble learning that has access to multi-core CPUs and GPGPU hardware in parallel or distributed environment. This should be considered as a major open research challenge that needs to be addressed for the complete adoption of the ML-based IDS models in the detection of real ICMPv6-based DDoS attacks.

B. SCALABILITY OF ML-BASED IDS IN LARGE SCALE NETWORKS

An individual ML-based IDS usually monitors the network traffic in a link-local IPv6 network to detect NDP-based DoS and DDoS attacks. However, increasing subnetworks naturally means an increase in the number of alarms generated by individual IDS in each subnet. As a consequence, there is an ever-increasing need for solutions that do not overwhelm human operators with unmanageable alarms. Therefore, a collaborative intrusion detection approach based on ensemble learning is especially useful, as it may ease the

discovery of NDP-based DoS and DDoS attacks by correlating information coming from various individual IDSs to improve the accuracy of a weak learner, thus can produce a lower number of false-positive alarms. This can also be considered as a significant research challenge for the detection of NDP-based DoS and DDoS attacks in large scale IPv6 networks.

C. EFFICIENT ANALYSIS OF FLOW-BASED NETWORK TRAFFIC

ML-based IDS models that use flow-based network traffic representation require preprocessing to generate network flow datasets before they are analyzed to predict ICMPv6-based DoS and DDoS attack patterns. For efficient analysis of flow-based network traffic, it should be interesting to try and implement a CIDS architecture consisting of preprocessing monitoring units to collect, generate, and share network flow dataset. Furthermore, they may also contain one or more analysis units to carry out the actual intrusion detection on the network flow dataset that is received from monitoring units to predict ICMPv6-based DoS and DDoS attack patterns.

D. LACK OF BENCHMARK DATASET

An intrusion detection dataset is a valuable tool for performance evaluation of the ML-based IDS model. The diversity in attack traffic is especially useful for benchmarking purposes to produce a more robust intrusion detection dataset. Although, Elejla *et. al* [44] prepared two publicly available intrusion detection datasets to evaluate the performance of ML-based IDS models using a flow-based network traffic representation. However, existing datasets lack DDoS attacks based on ICMPv6 error messages, which means that immediate attention needs to be paid to the evaluation and comparison of different ML-based IDS models that have used flow-based network traffic representation.

E. IMBALANCED DATASETS

Supervised learning-based IDS models, such as [59], [69], and [70], predicting anomalies as an attack or normal class needs attention to class imbalance factor, as learners would typically over-classify network data into one class if the majority of the training instances belongs to that class due to its increased probability existing in the dataset. A useful factor in terms of imbalanced datasets would be to concentrate on the proportion of classes in the confusion matrix. The MCC metric results may be more meaningful considering the proportion of classes within the confusion matrix. Therefore, there is a need to perform a comparative performance analysis of the proposed ML-based IDS models especially focusing on imbalanced datasets in the detection of ICMPv6-based DoS and DDoS attacks.

F. A SOLE PERFORMANCE EVALUATION METRIC WITH BINARY CLASS DATASET

DA as the only metric for performance evaluation of the ML-based IDS may mislead the researcher with a high percentage in the case of a binary class dataset, as in [62]. Classifier such as NB, that outputs negative class, may give 99% DA even though binary class dataset with a positive class contains 1% instances only. Therefore, this is a significant research issue that needs to be addressed by performing a comparative performance analysis of proposed ML-based IDS models with binary class datasets in the detection of ICMPv6-based DoS and DDoS attacks.

VI. NEW RESEARCH DIRECTIONS

A. LEVERAGING NEW NETWORK ARCHITECTURE

A new research direction can be combined with recent technological advancements aimed at leveraging a new network architecture, namely the Software-defined Networking (SDN), to propose ML-based IDS models for the ICMPv6-based DoS and DDoS attacks detection problem. SDN is an emerging architecture that decouples network management and routing functions so that network control is easily configurable [34]. The segregation of the control plane from the data plane makes it easier to efficiently manage the network and its protocols dynamically. Although SDN features can be helpful in defeating DDoS attacks, they may also be vulnerable to DDoS attacks due to the centralized characteristics and open programmable behavior of SDN [34]. In particular, a DDoS attack targets the SDN controller by sending a large number of malicious packets such that the entire network is compromised as a single point of failure. Conventional IDS solutions may not work well with large datasets, as in the case of DDoS attacks, so it is necessary to implement the IDS based on ML techniques for the detection of ICMPv6-based DoS and DDoS attacks [74].

In the literature, few IDS models have been proposed by researchers to detect ICMPv4-based DDoS attacks using ML techniques for SDN-based architectures, such as [75]–[77], and [78]. However, the common technique used by administrators as a solution for ICMPv4-based DDoS attacks was the rate limitation or complete blocking of ICMPv4 messages in an IPv4 network [31]. Additionally, unlike the ICMPv4 messages, the IPv6 implementation mandates the use of ICMPv6 messages for additional functions such as SLAAC, DAD, PMTUD, RD, and ND. Hence, existing ML-based IDS models may not be effective in detecting ICMPv6-based DoS and DDoS attacks due to new ICMPv6 vulnerability classes, and thus the detection of attacks in SDN-based architecture requires an efficient solution. Most importantly, ML-based IDS may be implemented in SDN controllers to improve classification results with high accuracy by extracting new features from ICMPv6-based DoS and DDoS attack cases, thus can serve as the SDN security solution. Therefore, the features of SDN encourages revolutionary implementations by dictating a new networking paradigm capable of implementing the ML-based IDS for the ICMPv6-based DoS and DDoS attacks detection problem.

B. ENHANCING THE SECURITY OF THE INTERNET OF THINGS (IOT) NETWORK

Recent research shows that IoT networks are facing difficulty in securing the overall availability of the IoT network with

rapid development in different areas [79]. In addition to affecting the security of IoT networks, the hazard posed by infested internet-connections threatens an overall internet environment which may leverage the vulnerable objects (smart devices) deployed as botnets to launch a DDoS attack [36]. There are many types of vulnerabilities in IoT networks, including DoS and DDoS vulnerabilities based on ICMPv6 messages, that need to be addressed in future research work. For instance, IPv6, as the network-layer protocol in the IoT architecture, provides the foundation for the interconnection of computing devices embedded in everyday objects. However, smart devices may be flooded with ICMPv6 Echo packets using spoofed IP addresses to launch a direct DDoS attack, or with a UDP-based DDoS reflection attack having a spoofed victim address as the source address [80].

In addition, IPv6 over Low Power Wireless Personal Area Networks (6LoWPAN) [81] is an adaptation layer protocol that defines optimization mechanisms such as compression, fragmentation, and reassembly to transport IP packets over IEEE 802.15.4 networks. In a normal scenario, the Routing Protocol for Low-Power and Lossy Networks (RPL) [82] constructs a Destination-Oriented Directed Acyclic Graph (DODAG) between the nodes in a 6LoWPAN network so that the routes in the DODAG are formed. However, RPL uses newly introduced ICMPv6 control messages from which an RPL-specific DODAG Information Solicitation (DIS) attack may be launched. For instance, a malicious RPL node can target other RPL nodes by sending an overwhelming amount of DIS messages, causing the recipient nodes to respond by sending the DODAG Information Option (DIO) messages. This will lead to congestion in the network and energy usage of resource-constrained nodes.

In the literature, researchers proposed ML-based IDS solutions using a modified version of benchmark datasets that are developed for non-IoT contexts, such as [83] and [84]. In addition, some researchers have proposed IDS solutions that target only RPL-based attacks in IoT networks, such as [85] and [86]. Specifically, there is an ever-increasing need for a public benchmark dataset in the real IoT context that can validate the selection of the new ICMPv6 features to test and verify the performance of different ML-based IDS solutions. Therefore, a new research direction can also be the enhancement of security in IoT networks against ICMPv6-based DoS and DDoS attacks using ML-based IDS solutions. Most notably, by deploying ML-based IDS in IoT with a focus on ICMPv6-based DoS and DDoS attack detection, it is possible to improve detection accuracy as well as to overcome the problem of traditional IoT IDS solutions suffering from high FPR.

C. EXTRACTION OF NEW FEATURES FOR ICMPV6-BASED DDOS ATTACK DETECTION IN FOG COMPUTING ENVIRONMENT

Fog computing was proposed to be integrated with IoT to extend cloud services to the edge of the network, bringing computing, communication, and storage closer to the edge devices and end-users, thus reducing the response time for end-user computing. However, the vulnerabilities of such devices remain a major concern, as availability is often threatened by external threats such as ICMPv6-based DDoS attacks. A new research direction may be the use of ML techniques to extract new features from ICMPv6 packet attributes in a fog computing environment to implement ML-based IDS solutions. Therefore, the detection of an ICMPv6-based DDoS attack may be handled at the edge of the network with the help of the ML-based IDS.

VII. BLOCKCHAIN APPLICABILITY

Blockchain is considered to be disruptive technology across multiple domains after the evolution of the theory introduced by Haber and Stornett [87]. The major contribution to blockchain's popularity was the introduction of the Bitcoin proposed by Nakamoto [88], which attracted much attention to this technology. This section provides an overview of blockchain technology and its application to ML-based IDS models for the detection of ICMPv6-based DoS and DDoS attacks.

A. BLOCKCHAIN PRIMITIVES

The blockchain consists of two main components: 1) a chain of blocks connected by a secure hash, thus storing the data in the blockchain at a particular moment in time; and 2) a network of nodes maintaining the state of the blockchain by following a consensus mechanism. The key contribution of the blockchain is the tamper-proof nature of the data stored in the blockchain, as changes in the state of the blockchain require the approval of the participants involved in the consensus mechanism.

1) SECURE HASH

All participants in the blockchain network maintain the state of the blockchain in a decentralized distributed environment, participating nodes compute a cryptographically secure hash (referred to as block hash) to add new blocks to the blockchain after successful consensus approval. Cryptographically secure hash computing requires a collision-free hash function that maps an arbitrary-length input to a fixed-length n-bit output as discussed in [89]. Meng *et al.* [90] report that the computation of a secure hash must satisfy the following requirements:

a) Each secure hash must be the result of some computational efforts to ensure pre-image resistance: Given a hash value h, it should require $\mathcal{O}(2^n)$ effort to compute an x, such that $H(x) = h$.
b) A different arbitrary-length input can only be mapped to the same hash provided that some computational efforts are made to find a secure hash to ensure another preimage resistance requirement: Given an input x and its hash value $h = H(x)$, it should require $\mathcal{O}(2^n)$ effort to compute an $x' \neq x$, such that $H(x') = h$.
c) The computational result of a secure hash is only the same provided that some computational efforts

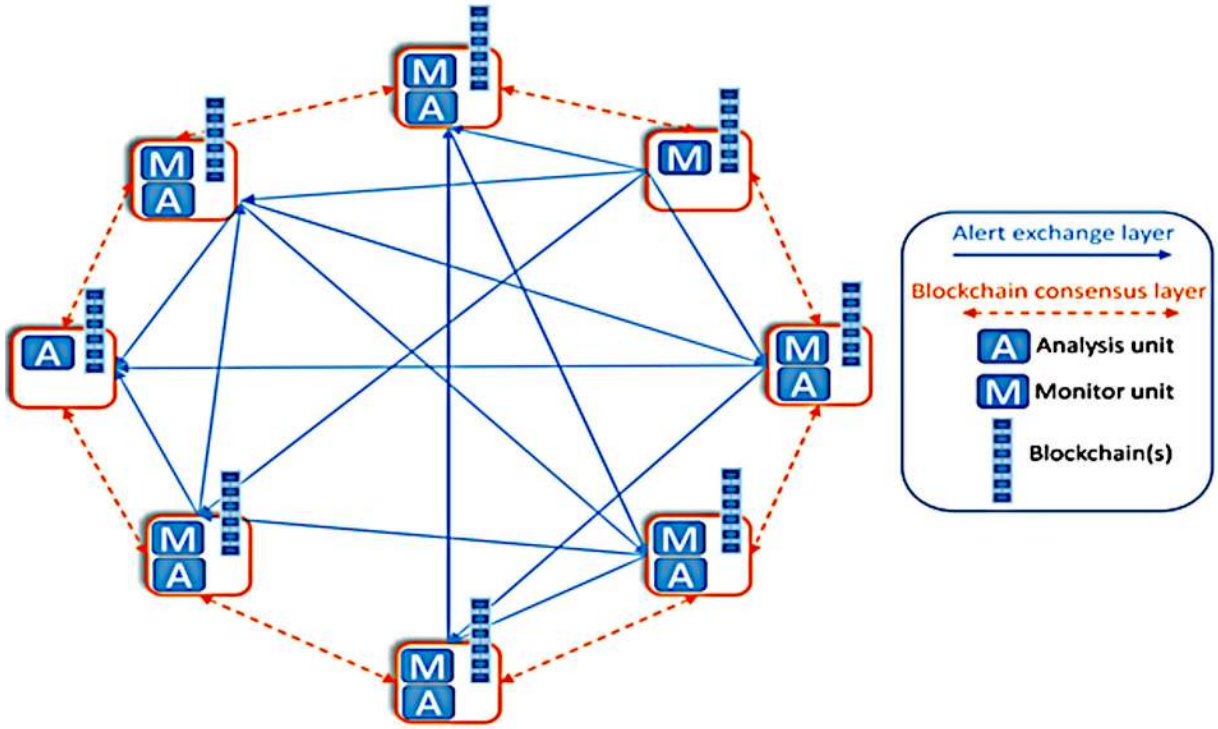

FIGURE 11. Blockchain-based CIDS architecture [91].

are made with different arbitrary-length inputs passed to the hash function to ensure collision resistance: It should require $\mathcal{O}\left(2^{n/2}\right)$ effort to compute any two $x \neq x'$ such that $H(x) = H(x')$.

2) CONSENSUS MECHANISM

The strength of the blockchain lies in the consensus mechanism that ensures a cryptographically secure mechanism for obtaining an open verifiable and immutable sequence of records (referred to as blocks) chronologically ordered by discrete timestamps. The blockchain follows a decentralized distributed environment where the state of the blockchain is generally shared and synchronized across all participants in a Peer-to-Peer (P2P) network.

The participation of all entities in a blockchain network is based on the type of blockchain and the permission granted to the participants either as readers or as writers. The public blockchain allows each participant on the blockchain network to access and reject or verify records based on a consensus mechanism. Whereas the types of private or consortium blockchain limit the accessibility of records. The consensus mechanism involves obtaining approval from the participating entities and blocks are appended to the blockchain in chronological order of their verification.

B. BLOCKCHAIN APPLICABILITY IN ICMPV6-BASED DOS AND DDOS ATTACKS DETECTION

The collaborative intrusion detection approach is particularly relevant for the detection of ICMPv6-based DDoS attacks. The collaborative architecture allows the sharing of information with other nodes on the network, either in the form of correlated alarms or in the form of attack detection models; thus minimizes the rate of false-positive alarms and is capable of detecting attacks that spread across entire subnetworks, such as NDP-based DoS attacks.

The collaborative architecture allows several monitoring units to collect and exchange network traffic data, and one or more analysis units carrying out the actual intrusion detection on the data obtained from monitoring units. Collaborative architecture based on ensemble learning may also be considered to improve the accuracy of weak learners in real-time as well as to minimize the rate of false-positive alarms. The CIDS architecture consisting of several analysis units is especially relevant due to the growing scale of the problem in real ICMPv6-based DDoS attack detection. Analysis units may work within a distributed framework to share the task of an attack detection model equally that is based on ensemble learning. In particular, each analyzing node should be capable of collecting, analyzing, sharing the output with other analyzing nodes, and then invoking alarms to implement the IDS model based on real-time ensemble learning.

The distributed framework uses the P2P network to share knowledge across multiple nodes that pose issues such as data sharing and trust computation among the nodes. On the one hand, the privacy concerns of an organization limit the sharing of their data, thus making it hard to optimize the detection and performance of the ML-based IDS model. On the other side, trust computation is another challenge to safeguard the integrity of the shared information among nodes by deterring insider attacks.

The inherent features of blockchain technology make it a viable solution to mitigate this problem, as it ensures data sharing and trust computation across multiple nodes by

applying a consensus mechanism that guarantees confidentiality, integrity, and availability in a distributed environment. As shown in Fig. 11, Alexopoulos *et al.* [91] have introduced the CIDS based on blockchain technology by considering raw alerts generated by individual IDS nodes as blockchain records. The trust computation among individual IDS nodes was achieved by maintaining the state of alerts generated by each IDS node, hence allowing IDS nodes to collect and exchange required information with each other. However, the authors focused only on the trust computation problem in CIDS architecture.

For the ICMPv6-based DoS and DDoS attack detection problem, blockchain applicability in CIDS architecture may be leveraged through a CIDS based on ensemble learning in a distributed environment. This would ease the discovery of ICMPv6-based DDoS attacks in real-time by sharing of knowledge across multiple analyzing nodes. Thus, the IDS would be more intelligent in decision-making in real-time which is required to tackle ICMPv6-based DDoS attacks. Further, It can also be more resilient to a single point of failure than current ML-based IDS solutions that use a centralized approach for network traffic capture, preprocessing, and analysis.

VIII. CONCLUSION

The main objective of the DoS and DDoS attacks is to halt the machine or the network, making it inaccessible to its intended users either by targeting the bandwidth of the network or by consuming the resources of the host. The ability of ML techniques to learn from examples has been a focal point for researchers, and the ML-based IDS is one of them for detecting DoS and DDoS attacks based on ICMPv6 messages.

We reviewed existing IDSs based on ML techniques that are proposed to detect ICMPv6-based DoS and DDoS attacks. We have learned that when traffic volumes grow by an order of magnitude, as in the case of ICMPv6-based DDoS attacks, existing IDSs based on single or hybrid classifiers will no longer be effective in real-time using sequential hardware. In addition, recent advanced technologies such as SDN, IoT, and fog computing encourage ML-based IDS solutions to be proposed in the context of ICMPv6-based DoS and DDoS attacks detection. Therefore, ML-based intrusion detection is still a significant research issue in the detection of ICMPv6-based DoS and DDoS attacks. Thus, this article concludes that the following issues could be useful for future research:

(i) Ensemble learning. Due to the growing scale of the problem by orders of magnitude in real ICMPv6-based DDoS attack detection, an IDS model based on the ensemble learning in a distributed environment is a promising area of research.

(ii) Collaborative architecture. Collaborative architecture can be useful, as correlating information from different individual classifiers would facilitate the discovery of ICMPv6-based DoS and DDoS attacks and would generally present a lower number of false-positive alarms.

(iii) Handling imbalanced datasets. Focusing on the proportion of the classes inside the confusion matrix would be a valuable consideration in the case of imbalanced datasets. The MCC metric results may be more meaningful as it considers the proportion of the classes inside the confusion matrix.

REFERENCES

[1] A. S. A. M. S. Ahmed, R. Hassan, and N. E. Othman, "IPv6 neighbor discovery protocol specifications, threats and countermeasures: A survey," *IEEE Access*, vol. 5, pp. 18187–18210, 2017.

[2] A. Conta, S. Deering, and M. Gupta, *Internet Control Message Protocol (ICMPv6)*, document RFC 4443, IETF, 2006.

[3] T. Narten, W. Simpson, E. Nordmark, and H. Soliman, *Neighbor Discovery for IP Version 6 (IPv6)*, document RFC 4861, IETF, 2007.

[4] A. K. Al-Ani, M. Anbar, A. Al-Ani, and D. R. Ibrahim, "Match-prevention technique against denial-of-service attack on address resolution and duplicate address detection processes in IPv6 link-local network," *IEEE Access*, vol. 8, pp. 27122–27138, 2020.

[5] R. M. A. Saad, M. Anbar, and S. Manickam, "Rule-based detection technique for ICMPv6 anomalous behaviour," *Neural Comput. Appl.*, vol. 30, no. 12, pp. 3815–3824, Dec. 2018.

[6] O. E. Elejla, M. Anbar, and B. Belaton, "ICMPv6-based DoS and DDoS attacks and defense mechanisms: Review," *IETE Tech. Rev.*, vol. 34, no. 4, pp. 390–407, Jul. 2017.

[7] O. E. Elejla, B. Belaton, M. Anbar, and A. Alnajjar, "Intrusion detection systems of ICMPv6-based DDoS attacks," *Neural Comput. Appl.*, vol. 30, no. 1, pp. 45–56, Jul. 2018.

[8] A. H. Bdair, R. Abdullah, S. Manickam, and A. K. Al-Ani, "Brief of intrusion detection systems in detecting ICMPv6 attacks," in *Proc. 6th Comput. Sci. Technol. (ICCST)*, Kota Kinabalu, Malaysia, vol. 603. Singapore: Springer, 2020, pp. 199–213.

[9] O. E. Elejla, B. Belaton, M. Anbar, B. Alabsi, and A. K. Al-Ani, "Comparison of classification algorithms on ICMPv6-based DDoS attacks detection," in *Computational Science and Technology* (Lecture Notes in Electrical Engineering), vol. 481. Singapore: Springer, 2019, pp. 347–357.

[10] N. K. Kasabov and Q. Song, "DENFIS: Dynamic evolving neural-fuzzy inference system and its application for time-series prediction," *IEEE Trans. Fuzzy Syst.*, vol. 10, no. 2, pp. 144–154, Apr. 2002.

[11] Google. (2020). *Percentage of Users that Access Google Over IPv6*. [Online]. Available: https://www.google.com/intl/en/ipv6/statistics.html

[12] J. B. Ard, "Internet protocol version six (IPv6) at UC Davis: Traffic analysis with a security perspective," M.S. thesis, Univ. California, Davis, CA, USA, 2012.

[13] R. C. Baishya and D. K. Bhattacharyya, "A complete detection and mitigation framework to protect a network from DDoS attacks," *IETE J. Res.*, pp. 1–18, Apr. 2019, doi: 10.1080/03772063.2019.1604173.

[14] A. K. Al-Ani, M. Anbar, S. Manickam, C. Y. Wey, Y.-B. Leau, and A. Al-Ani, "Detection and defense mechanisms on duplicate address detection process in IPv6 link-local network: A survey on limitations and requirements," *Arabian J. Sci. Eng.*, vol. 44, no. 4, pp. 3745–3763, Apr. 2019.

[15] Miniwatts Marketing Group. (Mar. 2020). *World Internet Usage and Population Statistics*. [Online]. Available: https://internetworldstats.com/stats.htm

[16] S. Deering and R. Hinden, *Internet Protocol, Version 6 (IPv6) Specification*, document RFC 2460, IETF, 1998.

[17] Internet Society. (Jun. 2012). *World IPv6 Launch*. [Online]. Available: https://www.worldipv6launch.org/

[18] S. Krishnan and S. Frankel, *IP Security (IPsec)*, document RFC 6071, IETF, 2011.

[19] J. Mirkovic and P. Reiher, "A taxonomy of DDoS attack and DDoS defense mechanisms," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 34, no. 2, pp. 39–53, Apr. 2004.

[20] O. Osanaiye, K.-K.-R. Choo, and M. Dlodlo, "Distributed denial of service (DDoS) resilience in cloud: Review and conceptual cloud DDoS mitigation framework," *J. Netw. Comput. Appl.*, vol. 67, pp. 147–165, May 2016.

[21] F. A. Barbhuiya, S. Biswas, and S. Nandi, "Detection of neighbor solicitation and advertisement spoofing in IPv6 neighbor discovery protocol," in *Proc. 4th Int. Conf. Secur. Inf. Netw. (SIN)*, New York, NY, USA, 2011, pp. 111–118.

[22] J. Gao and Y. Chen, "Detecting DOS/DDOS attacks under IPv6," in *Proc. Int. Conf. Cybern. Informat.*, vol. 163. New York, NY, USA: Springer, 2014, pp. 847–855.

[23] H. Scott and V. Erick, "Local network security," in *IPv6 Security* (Network Technology Series), 1st ed. Indianapolis, IN, USA: Pearson, 2008, pp. 183–193.

[24] Supriyanto, I. H. Hasbullah, R. K. Murugesan, and S. Ramadass, "Survey of Internet protocol version 6 link local communication security vulnerability and mitigation methods," *IETE Tech. Rev.*, vol. 30, no. 1, pp. 64–71, 2013.

[25] J. McCann, S. Deering, and J. Mogul, *Path MTU Discovery for IP Version 6*, document RFC 1981, IETF, 1996.

[26] N. C. Arjuman and S. Manickam, "A review on ICMPv6 vulnerabilities and its mitigation techniques: Classification and art," in *Proc. Int. Conf. Comput., Commun., Control Technol. (ICT)*, Kuching, Malaysia, Apr. 2015, pp. 323–327.

[27] C. E. Martin and J. H. Dunn, "Internet protocol version 6 (IPv6) protocol security assessment," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Orlando, FL, USA, Oct. 2007, pp. 1–7.

[28] R. Vida and L. Costa, Eds., *Multicast Listener Discovery Version 2 (MLDv2) for IPv6*, document RFC 3810, IETF, 2004.

[29] G. Bansal, N. Kumar, S. Nandi, and S. Biswas, "Detection of NDP based attacks using MLD," in *Proc. 5th Int. Conf. Secur. Inf. Netw. (SIN)*, New York, NY, USA, 2012, pp. 163–167.

[30] J. Ullrich, K. Krombholz, H. Hobel, A. Dabrowski, and E. Weippl, "IPv6 security: Attacks and countermeasures in a nutshell," in *Proc. 8th USENIX Conf. Offensive Technol. (WOOT)*. San Diego, CA, USA: USENIX Association, 2014, p. 5.

[31] M. Anbar, R. Abdullah, R. M. Saad, E. Alomari, and S. Alsaleem, "Review of security vulnerabilities in the IPv6 neighbor discovery protocol," in *Information Science and Applications* (Lecture Notes in Electrical Engineering), vol. 376. Singapore: Springer, 2016, pp. 603–612.

[32] M. Anbar, R. Abdullah, R. M. A. Saad, and I. H. Hasbullah, "Review of preventive security mechanisms for neighbour discovery protocol," *Adv. Sci. Lett.*, vol. 23, no. 11, pp. 11306–11310, Nov. 2017.

[33] C. Douligeris and A. Mitrokotsa, "DDoS attacks and defense mechanisms: Classification and state-of-the-art," *Comput. Netw.*, vol. 44, no. 5, pp. 643–666, Apr. 2004.

[34] R. Swami, M. Dave, and V. Ranga, "Software-defined networking-based DDoS defense mechanisms," *ACM Comput. Surv.*, vol. 52, no. 2, pp. 1–36, May 2019.

[35] J. Weber, "IPv6 security test laboratory," M.S. thesis, RUB, Bochum, Germany, 2013.

[36] N. Chaabouni, M. Mosbah, A. Zemmari, C. Sauvignac, and P. Faruki, "Network intrusion detection for IoT security based on learning techniques," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2671–2701, 3rd Quart., 2019.

[37] S. Rathore and J. H. Park, "Semi-supervised learning based distributed attack detection framework for IoT," *Appl. Soft Comput.*, vol. 72, pp. 79–89, Nov. 2018.

[38] E. Hodo, X. Bellekens, A. Hamilton, P.-L. Dubouilh, E. Iorkyase, C. Tachtatzis, and R. Atkinson, "Threat analysis of IoT networks using artificial neural network intrusion detection system," in *Proc. Int. Symp. Netw., Comput. Commun. (ISNCC)*, May 2016, pp. 1–6.

[39] F. Hosseinpour, P. V. Amoli, J. Plosila, T. Hämäläinen, and H. Tenhunen, "An intrusion detection system for fog computing and iot based logistic systems using a smart data approach," *Int. J. Digit. Content Technol. Appl.*, vol. 10, no. 5, pp. 34–46, 2016.

[40] N. Moustafa, B. Turnbull, and K.-K.-R. Choo, "An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of Internet of Things," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4815–4830, Jun. 2019.

[41] A. Verma and V. Ranga, "Machine learning based intrusion detection systems for IoT applications," *Wireless Pers. Commun.*, vol. 111, no. 4, pp. 2287–2310, Apr. 2020.

[42] C.-F. Tsai, Y.-F. Hsu, C.-Y. Lin, and W.-Y. Lin, "Intrusion detection by machine learning: A review," *Expert Syst. Appl.*, vol. 36, no. 10, pp. 11994–12000, 2009.

[43] F. Salo, M. Injadat, A. B. Nassif, A. Shami, and A. Essex, "Data mining techniques in intrusion detection systems: A systematic literature review," *IEEE Access*, vol. 6, pp. 56046–56058, 2018.

[44] O. E. Elejla, M. Anbar, B. Belaton, and S. Hamouda, "Labeled flow-based dataset of ICMPv6-based DDoS attacks," *Neural Comput. Appl.*, vol. 31, no. 8, pp. 3629–3646, Aug. 2019.

[45] S. Latif, Z. Zou, Z. Idrees, and J. Ahmad, "A novel attack detection scheme for the industrial Internet of Things using a lightweight random neural network," *IEEE Access*, vol. 8, pp. 89337–89350, 2020.

[46] H.-J. Liao, C.-H. R. Lin, Y.-C. Lin, and K.-Y. Tung, "Intrusion detection system: A comprehensive review," *J. Netw. Comput. Appl.*, vol. 36, no. 1, pp. 16–24, Jan. 2013.

[47] M. F. Umer, M. Sher, and Y. Bi, "Flow-based intrusion detection: Techniques and challenges," *Comput. Secur.*, vol. 70, pp. 238–254, Sep. 2017.

[48] L. Yao, L. ZhiTang, and L. Shuyu, "A fuzzy anomaly detection algorithm for IPv6," in *Proc. 2nd Int. Conf. Semantics, Knowl. Grid (SKG)*, Guilin, China, 2006, pp. 67–68.

[49] S. Forrest, B. Javornik, R. E. Smith, and A. S. Perelson, "Using genetic algorithms to explore pattern recognition in the immune system," *Evol. Comput.*, vol. 1, no. 3, pp. 191–211, Sep. 1993.

[50] R. M. A. Saad, "ICMPv6 flood attack detection using DENFIS algorithms," *Indian J. Sci. Technol.*, vol. 7, no. 2, pp. 168–173, Feb. 2013.

[51] Q. Song and N. Kasabov, "ECM-A novel on-line, evolving clustering method and its applications," in *Proc. 5th Biannual Conf. Artif. Neural Netw. Expert Syst.*, 2001, pp. 87–92.

[52] O. E. O. Elejla, "Flow-representation approach for ICMPV6-based DDOS attacks setection," Ph.D. dissertation, School Comput. Sci., Univ. Sci. Malaysia, Penang, Malaysia, 2018.

[53] A. A. Bahashwan, M. Anbar, and S. M. Hanshi, "Overview of IPv6 based DDoS and DoS attacks detection mechanisms," in *Proc. Int. Conf. Adv. Cyber Secur. (ACeS)*, in Communications in Computer and Information Science, Penang, Malaysia, vol. 1132. Singapore: Springer, 2020, pp. 153–167.

[54] A. Salih, X. Ma, and E. Peytchev, "Detection and classification of covert channels in IPv6 using enhanced machine learning," in *Proc. Int. Conf. Comput. Technol. Inf. Syst. (ICCTIS)*, Dubai, UAE, 2015, pp. 1–7.

[55] Z. Liu and Y. Lai, "A data mining framework for building intrusion detection models based on IPv6," in *Proc. Int. Conf. Inf. Secur. Assurance (ISA)*, in Lecture Notes in Computer Science, Seoul, South Korea, vol. 5576. Berlin, Germany: Springer, 2009, pp. 608–618.

[56] R. Agrawal and S. Ramkrishnan, "Fast algorithms for mining association rules," in *Proc. 20th Int. Conf. Very Large Data Bases (VLDB)*. Santiago de Chile, Chile: Morgan Kaufmann, 1994, pp. 478–499.

[57] M. Zulkiflee, M. Azmi, S. Ahmad, S. Sahib, and M. Ghani, "A framework of features selection for IPv6 network attacks detection," *WSEAS Trans. Commun.*, vol. 14, no. 46, pp. 399–408, 2015.

[58] A. Moraglio, C. Di Chio, and R. Poli, "Geometric particle swarm optimisation," in *Proc. Eur. Conf. Genet. Program. (EuroGP)*, in Lecture Notes in Computer Science, Valencia, Spain, vol. 4445. Berlin, Germany: Springer, 2007, pp. 125–136.

[59] B. Vrat, N. Aggarwal, and S. Venkatesan, "Anomaly detection in IPv4 and IPv6 networks using machine learning," in *Proc. Annu. IEEE India Conf. (INDICON)*, New Delhi, India, Dec. 2015, pp. 1–6.

[60] UCI Machine Learning Repository. (1999). *KDD Cup 1999 Data*. [Online]. Available: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

[61] R. M. A. Saad, M. Anbar, S. Manickam, and E. Alomari, "An intelligent ICMPv6 DDoS flooding-attack detection framework (v6IIDS) using back-propagation neural network," *IETE Tech. Rev.*, vol. 33, no. 3, pp. 244–255, May 2016.

[62] F. Najjar, M. M. Kadhum, and H. El-Taj, "Detecting neighbor discovery protocol-based flooding attack using machine learning techniques," in *Advances in Machine Learning and Signal Processing*, (Lecture Notes in Electrical Engineering), vol. 387. Cham, Switzerland: Springer, 2016, pp. 129–139.

[63] B. Sasha, "A strict anomaly detection model for IDS," Phrack Mag., St. Louis, MO, USA, Tech. Rep., 2000, vol. 38, no. 56. [Online]. Available: http://phrack.org/issues/56/11.html

[64] V. Hauser. (2020). *The Hackers Choice (THC) IPv6 Attack Toolkit*. [Online]. Available: https://github.com/vanhauser-thc/thc-ipv6

[65] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software," *ACM SIGKDD Explor. Newslett.*, vol. 11, no. 1, pp. 10–18, Nov. 2009.

[66] I. Ahmad, M. Basheri, M. J. Iqbal, and A. Rahim, "Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection," *IEEE Access*, vol. 6, pp. 33789–33795, 2018.
[67] M. Anbar, R. Abdullah, B. N. Al-Tamimi, and A. Hussain, "A machine learning approach to detect router advertisement flooding attacks in next-generation IPv6 networks," *Cognit. Comput.*, vol. 10, no. 2, pp. 201–214, Apr. 2018.
[68] M. Anbar, R. Abdullah, I. H. Hasbullah, Y.-W. Chong, and O. E. Elejla, "Comparative performance analysis of classification algorithms for intrusion detection system," in *Proc. 14th Annu. Conf. Privacy, Secur. Trust (PST)*, Auckland, New Zealand, 2016, pp. 282–288.
[69] S. Patel, A. Gupta, Nikhil, S. Kumari, M. Singh, and V. Sharma, "Network traffic classification analysis using machine learning algorithms," in *Proc. Int. Conf. Adv. Comput., Commun. Control Netw. (ICACCCN)*, Greater Noida, India, Oct. 2018, pp. 1182–1187.
[70] O. E. Elejla, M. Anbar, B. Belaton, and B. O. Alijla, "Flow-based IDS for ICMPv6-based DDoS attacks detection," *Arabian J. Sci. Eng.*, vol. 43, no. 12, pp. 7757–7775, Dec. 2018.
[71] A. A. Alsadhan, A. Hussain, and M. M. Alani, "Detecting NDP distributed denial of service attacks using machine learning algorithm based on flow-based representation," in *Proc. 11th Int. Conf. Develop. eSyst. Eng. (DeSE)*, Cambridge, U.K., Sep. 2018, pp. 134–140.
[72] A. Alsadhan, A. Hussain, P. Liatsis, M. Alani, H. Tawfik, P. Kendrick, and H. Francis, "Locally weighted classifiers for detection of neighbor discovery protocol distributed denial-of-service and replayed attacks," *Trans. Emerg. Telecommun. Technol.*, pp. 1–15, Jul. 2019, doi: 10.1002/ett.3700.
[73] P. Englert, "Locally weighted learning," *Seminar Class Auton. Learn. Syst.*, vol. 1, no. 1, pp. 1–9, 2012.
[74] R. Swami, M. Dave, and V. Ranga, "DDoS attacks and defense mechanisms using machine learning techniques for SDN," in *Security and Privacy Issues in Sensor Networks and IoT*. Hershey, PA, USA: IGI Global, 2020, pp. 193–214.
[75] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 785–794.
[76] J. Ye, X. Cheng, J. Zhu, L. Feng, and L. Song, "A DDoS attack detection method based on SVM in software defined network," *Secur. Commun. Netw.*, vol. 2018, pp. 1–8, Apr. 2018.
[77] Y. Yu, L. Guo, Y. Liu, J. Zheng, and Y. Zong, "An efficient SDN-based DDoS attack detection and rapid response platform in vehicular networks," *IEEE Access*, vol. 6, pp. 44570–44579, 2018.
[78] B. Han, X. Yang, Z. Sun, J. Huang, and J. Su, "OverWatch: A cross-plane DDoS attack defense framework with collaborative intelligence in SDN," *Secur. Commun. Netw.*, vol. 2018, pp. 1–15, Jan. 2018.
[79] S. M. Tahsien, H. Karimipour, and P. Spachos, "Machine learning based solutions for security of Internet of Things (IoT): A survey," *J. Netw. Comput. Appl.*, vol. 161, Jul. 2020, Art. no. 102630.
[80] M. Šimon and L. Huraj, "A study of DDoS reflection attack on Internet of Things in IPv4/IPv6 networks," in *Proc. Comput. Sci. Line Conf.* Cham, Switzerland: Springer, 2019, pp. 109–118.
[81] G. Montenegro, N. Kushalnagar, J. Hui, and D. Culler, *Transmission of IPv6 Packets Over IEEE 802.15. 4 Networks*, document RFC 4944, IETF, 2007.
[82] T. Winter, P. Thubert, A. Brandt, J. W. Hui, R. Kelsey, P. Levis, K. Pister, R. Struik, J.-P. Vasseur, and R. K. Alexander, *RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks*, document RFC 6550, IETF, 2012.
[83] J. Liu, B. Kantarci, and C. Adams, "Machine learning-driven intrusion detection for Contiki-NG-based IoT networks exposed to NSL-KDD dataset," in *Proc. 2nd ACM Workshop Wireless Secur. Mach. Learn.*, Jul. 2020, pp. 25–30.
[84] H. H. Pajouh, R. Javidan, R. Khayami, A. Dehghantanha, and K.-K.-R. Choo, "A two-layer dimension reduction and two-tier classification model for anomaly-based intrusion detection in IoT backbone networks," *IEEE Trans. Emerg. Topics Comput.*, vol. 7, no. 2, pp. 314–323, Apr. 2019.
[85] P. P. Ioulianou and V. G. Vassilakis, "Denial-of-service attacks and countermeasures in the RPL-based Internet of Things," in *Computer Security*. Cham, Switzerland: Springer, 2019, pp. 374–390.
[86] L. Wallgren, S. Raza, and T. Voigt, "Routing attacks and countermeasures in the RPL-based Internet of Things," *Int. J. Distrib. Sensor Netw.*, vol. 9, no. 8, Aug. 2013, Art. no. 794326.
[87] S. Haber and W. Stornetta, "How to time-stamp a digital document," *J. Cryptol.*, vol. 3, no. 2, pp. 99–111, 1991.
[88] S. Nakamoto. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*. [Online]. Available: https://bitcoin.org/ bitcoin.pdf
[89] I. B. Damgård, "Collision free hash functions and public key signature schemes," in *Advances in Cryptology* (Lecture Notes in Computer Science), Davos, Switzerland, vol. 304. Berlin, Germany: Springer, 1988, pp. 203–216.
[90] W. Meng, E. W. Tischhauser, Q. Wang, Y. Wang, and J. Han, "When intrusion detection meets blockchain technology: A review," *IEEE Access*, vol. 6, pp. 10179–10188, 2018.
[91] N. Alexopoulos, E. Vasilomanolakis, N. R. Ivánkó, and M. Mühlhäuser, "Towards blockchain-based collaborative intrusion detection systems," in *Critical Information Infrastructures Security* (Lecture Notes in Computer Science), Lucca, Italy, vol. 10707. Cham, Switzerland: Springer, 2018, pp. 107–118.

MOHAMMAD TAYYAB received the M.C.A. degree from Jamia Millia Islamia University, New Delhi, India, in 2008. He is currently pursuing the Ph.D. degree with the School of Computer Science, USM. For almost ten years, he has served as a Lecturer at the College of Computer Science, Jazan University, Saudi Arabia. His research interests include IDS, blockchain, and machine learning.

BAHARI BELATON received the B.App.Sc. (computer studies) degree from the South Australian Institute of Technology, Australia, the B.Sc. degree (Hons.) from Flinders University, and the Ph.D. degree from the University of Leeds, U.K. He is currently a Professor and the Dean of the School of Computer Science, USM. His research interests include scientific data visualization, computer graphics, and network security.

MOHAMMED ANBAR (Member, IEEE) received the Ph.D. degree in advanced computer networks from USM. He is currently a Senior Lecturer with the National Advanced IPv6 Centre (NAv6), USM. His current research interests include malware detection, web security, IDS, intrusion prevention systems (IPS), network monitoring, the Internet of Things (IoT), and IPv6 security.